

OHCO  
63-07-01  
#3

OIPE

## RAW SEQUENCE LISTING

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:56

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

4 <110> APPLICANT: Edwin L. Madison  
 5> Edgar O. Ong  
 6 Jiunn-Chern Yeh  
 7 Corvas International, Inc.  
 9 <120> TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
 10 TRANSMEMBRANE SERINE PROTEASES, THE ENCODED PROTEINS AND  
 11 METHODS BASED THEREON  
 13 <130> FILE REFERENCE: 24745-1607  
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/09/776,191  
 C--> 16 <141> CURRENT FILING DATE: 2001-02-02  
 18 <150> PRIOR APPLICATION NUMBER: 60/213,124  
 19 <151> PRIOR FILING DATE: 2000-06-22  
 21 <150> PRIOR APPLICATION NUMBER: 60/234,840  
 22 <151> PRIOR FILING DATE: 2000-06-22  
 24 <150> PRIOR APPLICATION NUMBER: 60/179,982  
 25 <151> PRIOR FILING DATE: 2000-02-03  
 27 <150> PRIOR APPLICATION NUMBER: 60/183,542  
 28 <151> PRIOR FILING DATE: 2000-02-18  
 30 <150> PRIOR APPLICATION NUMBER: 09/657,968  
 31 <151> PRIOR FILING DATE: 2000-02-08  
 33 <160> NUMBER OF SEQ ID NOS: 72  
 35 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 37 <210> SEQ ID NO: 1  
 38 <211> LENGTH: 3147  
 39 <212> TYPE: DNA  
 40 <213> ORGANISM: Homo Sapien  
 42 <220> FEATURE:  
 43 <223> OTHER INFORMATION: Nucleotide encoding MTSP1  
 45 <221> NAME/KEY: CDS  
 46 <222> LOCATION: (23)...(2589)  
 48 <300> PUBLICATION INFORMATION:  
 49 <301> AUTHORS: O'Brien, T.J. and Tanimoto, H.  
 50 <308> DATABASE ACCESSION NO: GenBank AR081724  
 W--> 51 <310> PATENT DOCUMENT NUMBER: US Pat 5972616  
 52 <311> PATENT FILING DATE: 1998-02-20  
 53 <312> PUBLICATION DATE: 1999-10-26  
 55 <400> SEQUENCE: 1  
 56 tcaagagcgg cctcggggta cc atg ggg agc gat cgg gcc cgc aag ggc gga 52  
 57 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly  
 58 1 5 10  
 60 ggg ggc ccg aag gac ttc ggc gcg gga ctc aag tac aac tcc cgg cac 100  
 61 Gly Gly Pro Lys Asp Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His  
 62 15 20 25  
 64 gag aaa gtg aat ggc ttg gag gaa ggc gtg gag ttc ctg cca gtc aac 148  
 65 Glu Lys Val Asn Gly Leu Glu Glu Gly Val Glu Phe Leu Pro Val Asn  
 66 30 35 40  
 68 aac gtc aag aag gtg gaa aag cat ggc ccg ggg cgc tgg gtg gtg ctg 196

ENTERED

see page 5

## RAW SEQUENCE LISTING

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:56

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

69	Asn	Val	Lys	Lys	Val	Glu	Lys	His	Gly	Pro	Gly	Arg	Trp	Val	Val	Leu	
70			45					50					55				
72	gca	gcc	gtg	ctg	atc	ggc	ctc	ctc	ttg	gtc	ttg	ctg	ggg	atc	ggc	ttc	244
73	Ala	Ala	Val	Leu	Ile	Gly	Leu	Leu	Leu	Val	Leu	Leu	Gly	Ile	Gly	Phe	
74		60					65					70					
76	ctg	gtg	tgg	cat	ttg	cag	tac	cgg	gac	gtg	cgt	gtc	cag	aag	gtc	ttc	292
77	Leu	Val	Trp	His	Leu	Gln	Tyr	Arg	Asp	Val	Arg	Val	Gln	Lys	Val	Phe	
78	75					80					85					90	
80	aat	ggc	tac	atg	agg	atc	aca	aat	gag	aat	ttt	gtg	gat	gcc	tac	gag	340
81	Asn	Gly	Tyr	Met	Arg	Ile	Thr	Asn	Glu	Asn	Phe	Val	Asp	Ala	Tyr	Glu	
82					95					100					105		
84	aac	tcc	aac	tcc	act	gag	ttt	gta	agc	ctg	gcc	agc	aag	gtg	aag	gac	388
85	Asn	Ser	Asn	Ser	Thr	Glu	Phe	Val	Ser	Leu	Ala	Ser	Lys	Val	Lys	Asp	
86			110						115				120				
88	gcg	ctg	aag	ctg	ctg	tac	agc	gga	gtc	cca	ttc	ctg	ggc	ccc	tac	cac	436
89	Ala	Leu	Lys	Leu	Leu	Tyr	Ser	Gly	Val	Pro	Phe	Leu	Gly	Pro	Tyr	His	
90			125					130					135				
92	aag	gag	tcg	gct	gtg	acg	gcc	ttc	agc	gag	ggc	agc	gtc	atc	gcc	tac	484
93	Lys	Glu	Ser	Ala	Val	Thr	Ala	Phe	Ser	Glu	Gly	Ser	Val	Ile	Ala	Tyr	
94		140					145					150					
96	tac	tgg	tct	gag	ttc	agc	atc	ccg	cag	cac	ctg	gtg	gag	gag	gcc	gag	532
97	Tyr	Trp	Ser	Glu	Phe	Ser	Ile	Pro	Gln	His	Leu	Val	Glu	Glu	Ala	Glu	
98	155					160					165					170	
100	cgc	gtc	atg	gcc	gag	gag	cgc	gta	gtc	atg	ctg	ccc	ccg	cgg	gcg	cgc	580
101	Arg	Val	Met	Ala	Glu	Glu	Arg	Val	Val	Met	Leu	Pro	Pro	Arg	Ala	Arg	
102				175						180					185		
104	tcc	ctg	aag	tcc	ttt	gtg	gtc	acc	tca	gtg	gct	ttc	ccc	acg	gac		628
105	Ser	Leu	Lys	Ser	Phe	Val	Val	Thr	Ser	Val	Val	Ala	Phe	Pro	Thr	Asp	
106				190						195				200			
108	tcc	aaa	aca	gta	cag	agg	acc	cag	gac	aac	agc	tgc	agc	ttt	ggc	ctg	676
109	Ser	Lys	Thr	Val	Gln	Arg	Thr	Gln	Asp	Asn	Ser	Cys	Ser	Phe	Gly	Leu	
110			205					210					215				
112	cac	gcc	cgc	ggt	gtg	gag	ctg	atg	cgc	ttc	acc	acg	ccc	ggc	ttc	cct	724
113	His	Ala	Arg	Gly	Val	Glu	Leu	Met	Arg	Phe	Thr	Thr	Pro	Gly	Phe	Pro	
114		220					225					230					
116	gac	agc	ccc	tac	ccc	gct	cat	gcc	cgc	tgc	cag	tgg	gcc	ctg	cgg	ggg	772
117	Asp	Ser	Pro	Tyr	Pro	Ala	His	Ala	Arg	Cys	Gln	Trp	Ala	Leu	Arg	Gly	
118	235					240					245				250		
120	gac	gcc	gac	tca	gtg	ctg	agc	ctc	acc	ttc	cgc	agc	ttt	gac	ctt	gcg	820
121	Asp	Ala	Asp	Ser	Val	Leu	Ser	Leu	Thr	Phe	Arg	Ser	Phe	Asp	Leu	Ala	
122					255					260				265			
124	tcc	tgc	gac	gag	cgc	ggc	agc	gac	ctg	gtg	acg	gtg	tac	aac	acc	ctg	868
125	Ser	Cys	Asp	Glu	Arg	Gly	Ser	Asp	Leu	Val	Thr	Val	Tyr	Asn	Thr	Leu	
126			270						275				280				
128	agc	ccc	atg	gag	ccc	cac	gcc	ctg	gtg	cag	ttg	tgt	ggc	acc	tac	cct	916
129	Ser	Pro	Met	Glu	Pro	His	Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	
130			285					290					295				
132	ccc	tcc	tac	aac	ctg	acc	ttc	cac	tcc	tcc	cag	aac	gtc	ctg	ctc	atc	964
133	Pro	Ser	Tyr	Asn	Leu	Thr	Phe	His	Ser	Ser	Gln	Asn	Val	Leu	Leu	Ile	

## RAW SEQUENCE LISTING

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:56

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

134	300	305	310	
136	aca ctg ata acc aac act gag cgg cgg cat ccc ggc ttt gag gcc acc	1012		
137	Thr Leu Ile Thr Asn Thr Glu Arg Arg His Pro Gly Phe Glu Ala Thr			
138	315	320	325	330
140	ttc ttc cag ctg cct agg atg agc agc tgt gga ggc cgc tta cgt aaa	1060		
141	Phe Phe Gln Leu Pro Arg Met Ser Ser Cys Gly Gly Arg Leu Arg Lys			
142		335	340	345
144	gcc cag ggg aca ttc aac agc ccc tac tac cca ggc cac tac cca ccc	1108		
145	Ala Gln Gly Thr Phe Asn Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro			
146		350	355	360
148	aac att gac tgc aca tgg aac att gag gtg ccc aac aac cag cat gtg	1156		
149	Asn Ile Asp Cys Thr Trp Asn Ile Glu Val Pro Asn Asn Gln His Val			
150		365	370	375
152	aag gtg agc ttc aaa ttc ttc tac ctg ctg gag ccc ggc gtg cct gcg	1204		
153	Lys Val Ser Phe Lys Phe Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala			
154		380	385	390
156	ggc acc tgc ccc aag gac tac gtg gag atc aat ggg gag aaa tac tgc	1252		
157	Gly Thr Cys Pro Lys Asp Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys			
158	395	400	405	410
160	gga gag agg tcc cag ttc gtc gtc acc agc aac agc aac aag atc aca	1300		
161	Gly Glu Arg Ser Gln Phe Val Val Thr Ser Asn Ser Asn Lys Ile Thr			
162		415	420	425
164	gtt cgc ttc cac tca gat cag tcc tac acc gac acc ggc ttc tta gct	1348		
165	Val Arg Phe His Ser Asp Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala			
166		430	435	440
168	gaa tac ctc tcc tac gac tcc agt gac cca tgc ccg ggg cag ttc acg	1396		
169	Glu Tyr Leu Ser Tyr Asp Ser Ser Asp Pro Cys Pro Gly Gln Phe Thr			
170		445	450	455
172	tgc cgc acg ggg cgg tgt atc cgg aag gag ctg cgc tgt gat ggc tgg	1444		
173	Cys Arg Thr Gly Arg Cys Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp			
174		460	465	470
176	gcc gac tgc acc gac cac agc gat gag ctc aac tgc agt tgc gac gcc	1492		
177	Ala Asp Cys Thr Asp His Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala			
178	475	480	485	490
180	ggc cac cag ttc acg tgc aag aac aag ttc tgc aag ccc ctc ttc tgg	1540		
181	Gly His Gln Phe Thr Cys Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp			
182		495	500	505
184	gtc tgc gac agt gtg aac gac tgc gga gac aac agc gac gag cag ggg	1588		
185	Val Cys Asp Ser Val Asn Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly			
186		510	515	520
188	tgc agt tgt ccg gcc cag acc ttc agg tgt tcc aat ggg aag tgc ctc	1636		
189	Cys Ser Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu			
190		525	530	535
192	tcg aaa agc cag cag tgc aat ggg aag gac gac tgt ggg gac ggg tcc	1684		
193	Ser Lys Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser			
194		540	545	550
196	gac gag gcc tcc tgc ccc aag gtg aac gtc gtc act tgt acc aaa cac	1732		
197	Asp Glu Ala Ser Cys Pro Lys Val Asn Val Val Thr Cys Thr Lys His			
198	555	560	565	570

## RAW SEQUENCE LISTING

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:57

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

```

200 acc tac cgc tgc ctc aat ggg ctc tgc ttg agc aag ggc aac cct gag      1780
201 Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu
202          575          580          585
204 tgt gac ggg aag gag gac tgt agc gac ggc tca gat gag aag gac tgc      1828
205 Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp Glu Lys Asp Cys
206          590          595          600
208 gac tgt ggg ctg cgg tca ttc acg aga cag gct cgt gtt gtt ggg ggc      1876
209 Asp Cys Gly Leu Arg Ser Phe Thr Arg Gln Ala Arg Val Val Gly Gly
210          605          610          615
212 acg gat gcg gat gag ggc gag tgg ccc tgg cag gta agc ctg cat gct      1924
213 Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln Val Ser Leu His Ala
214          620          625          630
216 ctg ggc cag ggc cac atc tgc ggt gct tcc ctc atc tct ccc aac tgg      1972
217 Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp
218 635          640          645          650
220 ctg gtc tct gcc gca cac tgc tac atc gat gac aga gga ttc agg tac      2020
221 Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp Arg Gly Phe Arg Tyr
222          655          660          665
224 tca gac ccc acg cag tgg acg gcc ttc ctg ggc ttg cac gac cag agc      2068
225 Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly Leu His Asp Gln Ser
226          670          675          680
228 cag cgc agc gcc cct ggg gtg cag gag cgc agg ctc aag cgc atc atc      2116
229 Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg Leu Lys Arg Ile Ile
230          685          690          695
232 tcc cac ccc ttc ttc aat gac ttc acc ttc gac tat gac atc gcg ctg      2164
233 Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu
234          700          705          710
236 ctg gag ctg gag aaa ccg gca gag tac agc tcc atg gtg cgg ccc atc      2212
237 Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser Met Val Arg Pro Ile
238 715          720          725          730
240 tgc ctg ccg gac gcc tcc cat gtc ttc cct gcc ggc aag gcc atc tgg      2260
241 Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala Gly Lys Ala Ile Trp
242          735          740          745
244 gtc acg ggc tgg gga cac acc cag tat gga ggc act ggc gcg ctg atc      2308
245 Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile
246          750          755          760
248 ctg caa aag ggt gag atc cgc gtc atc aac cag acc acc tgc gag aac      2356
249 Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln Thr Thr Cys Glu Asn
250          765          770          775
252 ctc ctg ccg cag cag atc acg ccg cgc atg atg tgc gtg ggc ttc ctc      2404
253 Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met Cys Val Gly Phe Leu
254          780          785          790
256 agc ggc ggc gtg gac tcc tgc cag ggt gat tcc ggg gga ccc ctg tcc      2452
257 Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser
258 795          800          805          810
260 agc gtg gag gcg gat ggg cgg atc ttc cag gcc ggt gtg gtg agc tgg      2500
261 Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly Val Val Ser Trp
262          815          820          825
264 gga gac ggc tgc gct cag agg aac aag cca ggc gtg tac aca agg ctc      2548

```

## RAW SEQUENCE LISTING

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:57

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

```

265 Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu
266      830      835      840
268 cct ctg ttt cgg gac tgg atc aaa gag aac act ggg gta ta ggggccgggg 2599
269 Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly Val
270      845      850      855
272 ccacccaaat gtgtacacct gcggggccac ccacgttcca cccagtggtg caccgctgca 2659
273 ggctggagac tggaccgctg actgcaccag cgcgcccgaga acatacactg tgaactcaat 2719
274 ctccagggct ccaaatctgc ctagaaaacc tctcgcttcc tcagcctcca aagtggagct 2779
275 gggaggtaga aggggaggac actggtggtt ctactgaccc aactgggggc aaaggtttga 2839
276 agacacagcc tcccccgcca gccccaagct gggccgaggc gcgtttgtgt atatctgcct 2899
277 cccctgtctg taaggagcag cgggaacgga gcttcggagc ctccctcagt aagtggtggtg 2959
278 ggctgccgga tctgggctgt ggggcccttg ggccacgctc ttgaggaagc ccaggtcgg 3019
279 aggaccctgg aaaacagacg ggtctgagac tgaaattggt ttaccagctc ccaggtgga 3079
280 cttcagtggt tgtatttggt taaatgggta aaacaattta tttcttttta aaaaaaaaaa 3139
281 aaaaaaaaaa 3147
283 <210> SEQ ID NO: 2
284 <211> LENGTH: 855
285 <212> TYPE: PRT
286 <213> ORGANISM: Homo Sapien
288 <400> SEQUENCE: 2
289 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Gly Pro Lys Asp Phe
290 1 5 10 15
291 Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn Gly Leu
292 20 25 30
293 Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys Lys Val Glu
294 35 40 45
295 Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala Val Leu Ile Gly
296 50 55 60
297 Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln
298 65 70 75 80
299 Tyr Arg Asp Val Arg Val Gln Lys Val Phe Asn Gly Tyr Met Arg Ile
300 85 90 95
301 Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu Asn Ser Asn Ser Thr Glu
302 100 105 110
303 Phe Val Ser Leu Ala Ser Lys Val Lys Asp Ala Leu Lys Leu Leu Tyr
304 115 120 125
305 Ser Gly Val Pro Phe Leu Gly Pro Tyr His Lys Glu Ser Ala Val Thr
306 130 135 140
307 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser
308 145 150 155 160
309 Ile Pro Gln His Leu Val Glu Glu Ala Glu Arg Val Met Ala Glu Glu
310 165 170 175
311 Arg Val Val Met Leu Pro Pro Arg Ala Arg Ser Leu Lys Ser Phe Val
312 180 185 190
313 Val Thr Ser Val Val Ala Phe Pro Thr Asp Ser Lys Thr Val Gln Arg
314 195 200 205
315 Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu His Ala Arg Gly Val Glu
316 210 215 220
317 Leu Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala

```

Fy11.

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

## VERIFICATION SUMMARY

DATE: 02/28/2001

PATENT APPLICATION: US/09/776,191

TIME: 15:35:58

Input Set : A:\1607seq.001

Output Set: N:\CRF3\02282001\I776191.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:51 M:256 W: Invalid Numeric Header Field, Wrong PATENT DOCUMENT NUMBER:US NN/NNN,NNN  
L:1539 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13  
L:1554 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14